8

به نام خدا

دانشگاه تهران

پردیس دانشکده های فنی

دانشکده مهندسی برق و کامپیوتر

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

تمرین شماره‌ی 1

بخش عملی

**استاد درس:**

دکتر فدائی

دکتر یعقوب‌زاده

**نگارش:**

فاطمه محمدی 810199489

[اهداف: 3](#_Toc162646227)

[مقدمه: 3](#_Toc162646228)

[تعریف پروژه: 3](#_Toc162646229)

[داده‌ها و پارامتر‌های مسئله: 3](#_Toc162646230)

[پیاده سازی گام به گام پروژه: 3](#_Toc162646231)

[بخش صفر: مروری بر ورودی ها، کتابخانه ها و داده های ضروری: 3](#_Toc162646232)

[تعریف کتابخانه‌های مورد نیاز: 3](#_Toc162646233)

[تعریف ورودی‌ها: 3](#_Toc162646234)

[بخش اول: پیش پردازش و استخراج ویژگی: 4](#_Toc162646235)

[پیش پردازش: 4](#_Toc162646236)

[استخراج ویژگی: 4](#_Toc162646237)

[بخش دوم: آشنایی با HMM: 5](#_Toc162646238)

[بخش سوم: پیاده سازی مسئله: 5](#_Toc162646239)

[نتیجه گیری نهایی: 5](#_Toc162646240)

[راهکار برای توسعه و بهبود پروژه: 5](#_Toc162646241)

[سوالات: 5](#_Toc162646242)

[منابع استفاده شده: 5](#_Toc162646243)

بهینه سازی "کوله‌پشتی"

الگوریتم‌های ژنتیک

# اهداف:

توسعه یک سیستم تشخیص گفتار اعداد براساس HMM میباشد.

# مقدمه:

Hidden Markov Models (HMM) یکی از ابزارهای مهم در زمینه پردازش سیگنال های زمانی و تشخیص الگو هستند. در این مدل ها هر حالت به یک وضعیت مشخص مرتبط می شود و احتمال حرکت از یک وضعیت به وضعیت دیگر با توجه به وضعیت فعلی تعیین می شود. از این رو، HMM به عنوان یک ابزار قدتمند برای مدلسازی سیستم های پویا و تصمیم گیری در شرایط عدم قطعیت به شمار می آید. در زمینه تشخیص گفتار، HMM معمولاً برای مدل سازی دقیق تر وضعیت های مختلف گفتار، مانند حروف، کلمات یا فریم های زمانی کوتاه از گفتار، استفاده می شود.

# تعریف پروژه:

همانطور که اشاره شد هدف اصلی این پروژه توسعه یک سیستم تشخیص گفتار اعداد براساس HMM میباشد. برای اینکار مجموعه داده صوتی ای استفاده میشود و باید مراحل پیش پردازش، استخراج ویژگی، پیاده سازی مدل و تحلیل آن در طی پروژه انجام داد.

# داده‌ها و پارامتر‌های مسئله:

در این پروژه که یک پروژه داده محور است، مجموعه داده‌ای مورد نیاز، شامل ضبط های گفتاری از 6 گوینده مختلف است که هر کدام 50 بار ارقام 0 تا 9 را به زبان انگلیسی تلفظ می کنند. نام هر فایل صوتی به صورت {digitLabel}\_{speakerName}\_{index}.wav است که به ترتیب عدد گفته شده، نام گوینده و شماره نمونه را نشان میدهد.

# پیاده سازی گام به گام پروژه:

در ادامه به توضیح هر گام از پروژه به همراه علت انجام آن و تحلیل خروجی توابع صدا زده در هر گام میپردازیم.

## بخش صفر: مروری بر ورودی ها، کتابخانه ها و داده های ضروری:

### تعریف کتابخانه‌های مورد نیاز:

در این پروژه به چهار کتابخانه مهم نیاز داریم:

در ادامه علت استفاده از هریک از کتابخانه ها را بیان میکنیم.

### تعریف ورودی‌ها:

## بخش اول: پیش پردازش و استخراج ویژگی:

### پیش پردازش:

پیش پردازش داده ها اولین گام در هر پروژه داده محور است. هدف از پیش پردازش داده ها، بهبود کیفیت داده ها و آماده سازی آنها برای آموزش و آزمایش سیستم تشخیص گفتار است. مراحل معمول پیش پردازش داده ها می تواند شامل افزایش داده، حذف نویز، تقویت سیگنال، نرمال سازی و قطعه بندی باشد.

#### پرسش 1:

آیا قطعه بندی داده ها برای این دیتاسن مفید است؟ چرا؟

پاسخ:

### استخراج ویژگی:

استخراج و انتخاب ویژگی از مهم ترین مراحل هر پروژه هوشمصنوعی و یادگیری ماشین می باشند. استفاده از ویژگی مناسب می تواند تاثیر بالایی در خروجی مدل نهایی داشته باشد.

همچنین ویژگی های بسیار متفاوتی را می توان از یک محتوای صوتی استخراج کرد که هر کدام اطلاعات گوناگونی را به ما می دهند. از جمله این ویژگی ها می توان mel-spectrogram، Zero Crossing Rate، mfcc و chroma features را نام برد.

#### پرسش 2:

در مورد هر کدام از این ویژگی ها تحقیق کنید و روابط بین آن ها را توضیح دهید.

پاسخ:

در این پروژه از ضرایب Mel Frequency Cepstral Coefficient (MFCC) برای دسته بندی record ها استفاده میکنیم.

ضرایب MFCC به طور گسترده در زمینه های مرتبط با دسته بندی موسیقی و تشخیص گفتار استفاده شده اند.

در ادامه مجموعه ضرایب MFCC را برای هر کدام از نمونه های داده شده استخراج میکنیم.

پس از استخراج ویژگی ها، نمودار Heat Map مربوط به به ضرایب برای یک نمونه از هرکدام از دسته ها را رسم میکنیم.

#### پرسش 3:

پاسخ:

#### پرسش 4:

پاسخ:

#### پرسش 5:

پاسخ:

#### پرسش 6:

پاسخ:

## بخش دوم: آشنایی با HMM:

#### پرسش 1:

پاسخ:

#### پرسش 2:

پاسخ:

#### پرسش 3:

پاسخ:

#### پرسش 4:

پاسخ:

#### پرسش 5:

پاسخ:

## بخش سوم: پیاده سازی مسئله:

### پیاده سازی به کمک کتابخانه ها:

### پیاده سازی به کمک Scratch:

# نتیجه گیری نهایی:

# راهکار برای توسعه و بهبود پروژه:

# سوالات:

# منابع استفاده شده:

* Stackoverflow